



文章编号:1001-9081(2018)10-2918-05

DOI:10.11772/j.issn.1001-9081.2018040888

基于差分隐私保护的谱聚类算法

郑孝遥^{1,2*}, 陈冬梅^{1,2}, 刘雨晴^{1,2}, 尤 浩^{1,2}, 汪祥舜^{1,2}, 孙丽萍^{1,2}

(1. 安徽师范大学 计算机与信息学院, 安徽 芜湖 241002; 2. 网络与信息安全安徽省重点实验室(安徽师范大学), 安徽 芜湖 241002)
(*通信作者电子邮箱 zxiaoayao_2000@163.com)

摘要:针对传统的聚类算法存在隐私泄露的风险,提出一种基于差分隐私保护的谱聚类算法。该算法基于差分隐私模型,利用累计分布函数生成满足拉普拉斯分布的随机噪声,将该噪声添加到经过谱聚类算法计算的样本相似度的函数中,干扰样本个体之间的权重值,实现样本个体间的信息隐藏以达到隐私保护的目的。通过UCI数据集上的仿真实验,表明该算法能够在一定的信息损失度范围内实现有效的数据聚类,也可以对聚类数据进行保护。

关键词:差分隐私;谱聚类;敏感数据;隐私泄露

中图分类号:TP309.2 **文献标志码:**A

Spectral clustering algorithm based on differential privacy protection

ZHENG Xiaoyao^{1,2*}, CHEN Dongmei^{1,2}, LIU Yuqing^{1,2}, YOU Hao^{1,2}, WANG Xiangshun^{1,2}, SUN Liping^{1,2}

(1. School Computer and Information, Anhui Normal University, Wuhu Anhui 241002, China;
2. Anhui Provincial Key Laboratory of Network and Information Security (Anhui Normal University), Wuhu Anhui 241002, China)

Abstract: Aiming at the problem of privacy leakage in the application of traditional clustering algorithm, a spectral clustering algorithm based on differential privacy protection was proposed. Based on the differential privacy model, the cumulative distribution function was used to generate random noise that satisfies Laplace distribution. Then the noise was added to the sample similarity function calculated by the spectral clustering algorithm, which disturbed the weight values between the individual samples and realized information hiding between sample individuals for privacy protection. Experimental results of UCI dataset verify that the proposed algorithm can achieve effective data clustering within a certain degree of information loss, and can also protect clustered data.

Key words: differential privacy; spectral clustering; sensitive data; privacy leakage

0 引言

近年来,随着互联网与信息技术的蓬勃发展,海量数据的产生可以为研究者们提供许多有效的信息资源,对这些海量数据进行挖掘分析可以得到非常有价值的信息,其中聚类分析是有效手段之一,但是在聚类的过程中也存在着隐私泄露的风险。

目前,国内外研究者围绕隐私安全保护问题做了大量的工作,从已有的隐私保护技术来看: k -匿名及其扩展的保护模型技术已被广泛使用。该方法通过存储至少 k 条记录来隐藏一条数据记录从而达到隐私保护的目的,但是在攻击者掌握特定背景知识的情况下,它就存在不能抵抗一致性攻击的可能。为了克服这一缺点,研究者们不断尝试加以改进而出现了 l -多样性^[1]、 t -邻近^[2]、 (a,k) -匿名^[3]、泛化和随机化^[4]等隐私保护技术。

如今关于聚类分析在隐私保护方面的应用越来越多,而且聚类作为数据挖掘和机器学习的主要技术之一被广大学者

所研究,聚类分析中常用的隐私保护技术有数值扰动、数值旋转、数值匿名等方法。Oliverira等^[5]在2004年提出一种基于新的空间数据转换方法RT(Rotation-based Transformation),该方法的基础是基于几何的旋转变化来隐藏信息,并能保证旋转前后的属性依旧具有有效性;其缺点在于只能对低维度的数据进行变换,维度较高时会损耗数据的真实信息,计算量也较大,并且难以避免一致性攻击造成的隐私泄露。Mukherjee等^[6]在2006年针对文献[5]算法不易推广到其他应用场景中的问题而提出了一种新的采用傅里叶变换的数据扰动方法,其优点是在实现敏感信息隐藏的同时能够保持数据值的有效性,并保证欧几里得集中式和分布式场景中的距离精度很高。目前的研究在聚类的基础上加上差分隐私技术的研究还不够成熟,最早提出这一想法的是2005年Blum等^[7]给出的利用差分隐私和 k -means相结合的算法,使用管理员的身份将噪声引入到查询响应中来维护单个数据库条目的隐私。随后Dwork等^[8]在2010年对基于差分隐私的 k -means算法进行了更进一步的分析和改进,提出了一种可以计算查询函

收稿日期:2018-04-28;修回日期:2018-07-12;录用日期:2018-07-16。

基金项目:国家自然科学基金资助项目(61772034, 61602009);安徽省自然科学基金资助项目(1808085MF172)。

作者简介:郑孝遥(1981—),男,安徽芜湖人,副教授,博士研究生,CCF会员,主要研究方向:信息安全、个性化推荐; 陈冬梅(1994—),女,安徽天长人,硕士研究生,主要研究方向:信息安全、智能计算; 刘雨晴(1994—),女,安徽阜阳人,硕士研究生,主要研究方向:信息安全、数据挖掘; 尤浩(1997—),男,安徽颍上人,主要研究方向:信息安全、个性化推荐; 汪祥舜(1992—),男,安徽安庆人,硕士研究生,主要研究方向:个性化推荐; 孙丽萍(1980—),女,安徽芜湖人,教授,博士,CCF会员,主要研究方向:空间数据处理、智能计算。



数和查询序列敏感度的算法。国内的李杨等^[9]提出了 ϵ -差分隐私保护下的IDP k -means方法,该方法可以应对攻击者具有任意背景知识的攻击,并且很好地改善了数据聚类后的可用性问题,但是此方法不能解决分布式环境下的隐私泄露安全问题。因此,在2016年李洪成等^[10]提出了在分布式环境下满足 ϵ -差分隐私的 k -means算法,解决了传统隐私保护模型无法在分布式环境中应对任意背景知识攻击的问题。吴伟民等^[11]基于文献[9]提出了满足 ϵ -差分隐私保护的DP-DBScan算法,是应对传统的DBScan聚类算法存在的隐私泄露问题而研究出来的,该算法的实验结果表明与传统的DBScan聚类算法相比,加入拉普拉斯噪声的DP-DBScan算法能保持数据的有效性并实现隐私保护。刘晓迁等^[12]基于聚类的匿名化技术,并利用差分隐私保护模型对发布数据进行保护,该方法通过聚类实现匿名化,对匿名分化的数据添加随机噪声来扰动数据的真实值以实现隐私保护,并提高数据的可用性。目前,针对聚类方法结合差分隐私的研究还比较少,本文主要针对谱聚类算法,开展基于差分隐私保护的谱聚类算法研究。

本文主要针对聚类算法存在隐私泄露的问题,利用谱聚类算法将社交关系图网络化,使得具有社交关系的聚类在一起,为了防止这种社交关系泄露,在通过相似性函数计算权重时加上差分隐私的拉普拉斯随机噪声来进行扰动达到隐私保护的作用。

1 相关工作

1.1 谱聚类算法

目前,谱聚类算法在国内得到广泛的应用和研究,最常应用的领域有机器学习、大数据挖掘、图像分割、文本挖掘等^[13]。该算法的流程简单,并且与传统的 k -means和最大期望(Expectation Maximization, EM)算法相比适用性更强,在任何的数据样本空间中都会聚类出最好的聚类效果。谱聚类的主要思想是基于图谱理论的分割技术,该算法将样本数据看成无向带权图中的一个个顶点,然后利用相似性函数计算出各顶点之间的值,权重值代表各个顶点间的相似度,然后采用图谱的分割准则将数据分割聚类。

一般谱聚类算法的相似性函数采用余弦函数和高斯函数,具体定义如下:

$$\text{sim}(X, Y) = \cos \theta = \frac{\mathbf{X} \cdot \mathbf{Y}}{\|\mathbf{X}\| \cdot \|\mathbf{Y}\|} \quad (1)$$

$$W_{ij} = \exp\left(-\frac{d(s_i, s_j)}{2\sigma^2}\right) \quad (2)$$

图分割算法准则主要有以下几个准则:

1) 图分割最小分割法准则:

$$\text{Cut}(A, B) = \sum_{u \in A, v \in B} w(u, v) \quad (3)$$

2) 规范化分割准则:

$$\text{Normalized_Cut}(A, B) = \frac{\text{Cut}(A, B)}{\text{sum}(A, V)} + \frac{\text{Cut}(B, A)}{\text{sum}(B, V)} \quad (4)$$

3) 最小最大分割准则:

$$\text{Mcut}(A, B) = \frac{\text{Cut}(A, B)}{\text{sum}(A, A)} + \frac{\text{Cut}(B, B)}{\text{sum}(B, B)} \quad (5)$$

4) 比例分割准则:

$$Rcut(A, B) = \frac{\text{Cut}(A, B)}{\min(|A|, |B|)} \quad (6)$$

谱聚类算法的主要算法流程如算法1所示。

算法1 谱聚类算法。

输入 n 个数据点集,聚类数目 k 。

输出 得到 k 个簇划分。

a) 通过高斯核函数的距离公式计算相似性矩阵;

b) 利用相似矩阵构建邻接矩阵 N 和度矩阵 G ;

c) 由第 b) 步得到的邻接矩阵和度矩阵求出拉普拉斯矩阵 $L = G^{1/2}NG^{1/2}$;

d) 得到拉普拉斯矩阵后选取前 k 个最大特征值对应的特征向量;

e) 将特征向量标准化,然后将样本数据点映射到基于一个或多个确定的降维空间中去;

f) 基于新的数据点空间,将特征矩阵的每一行看成一个样本点利用 k -means 将它聚为 k 类。

1.2 差分隐私

差分隐私保护模型具有严格的数学理论基础,该模型的基本思想是对数据添加噪声来达到隐私保护的目的^[14]。该保护方法不需要关心攻击者强大计算能力和任何的背景知识,即使攻击者拥有除一条记录以外的所有数据记录也能保证这条记录的敏感信息不会被披露,这种机制能够对样本数据中个体敏感信息进行特定的保护,而且还不引起数据分布的变化。差分隐私的具体定义如下:

定义1 假设数据集 D 和 D' 是相邻数据集,两个数据集完全相同或至多相差一条数据记录,给定一个随机算法 S , $\text{Range}(S)$ 是算法 S 的取值范围, R_M 是数据集上的输出结果, $\Pr[X]$ 是事件 X 的披露风险,则算法 S 满足 ϵ -差分隐私的保护模型定义:

$$\Pr[S(D) = R_M] \leq \exp(\epsilon) \Pr[S(D') = R_M] \quad (7)$$

隐私披露风险的值由随机算法 S 控制,通过限制 ϵ 的大小来控制隐私保护的安全度: ϵ 越小引入的随机噪声越大,隐私保护安全性越高; ϵ 越大引入的随机噪声越小,隐私保护安全性越低。

定义2 对于函数 $F: D \rightarrow \mathbf{R}^d$ 的敏感度定义如下:

$$\Delta F = \max_{D, D'} \|F(D) - F(D')\|_1 \quad (8)$$

其中: $\|\cdot\|_1$ 表示一阶范数; F 是查询函数; d 是记录数据的属性维度; D 和 D' 是至多相差一条数据记录的数据集; \mathbf{R} 表示的实数空间。

由定义1可以看出随机函数的选择与攻击者的背景知识无关,所以任意一条记录的增加或者删除都不会影响查询结果的输出。该定义从理论上满足了差分隐私对隐私披露风险的要求,而具体的实现还是依靠添加噪声机制来实现。

1.3 两种噪声机制

实现差分隐私保护的噪声添加机制有两种:一种是基于数值型的拉普拉斯噪声机制;另一种是针对非数值型数据的指数机制^[15]。

1.3.1 拉普拉斯噪声机制

文献[16]中提出了对于数值型的数据采用拉普拉斯机制对数据的真实值进行扰动来达到隐私安全保护。其定义如



下：

假设噪声 x 服从尺度为 $b = \Delta F/\varepsilon$ 的拉普拉斯分布，则噪声函数的分布是标准差为 $\sqrt{2}b$ 的对称指数分布，函数表达形式如下：

$$\text{Laplace}(b) = \exp(-|x|/b) \quad (9)$$

其概率密度函数表达形式是：

$$p(x) = \exp(-|x|/b)/(2b) \quad (10)$$

其累积分布函数的表达形式是：

$$D(x) = [1 + \text{sgn}(x) \times (1 - \exp(-|x|/b))] / 2 \quad (11)$$

其概率密度函数如图1所示。

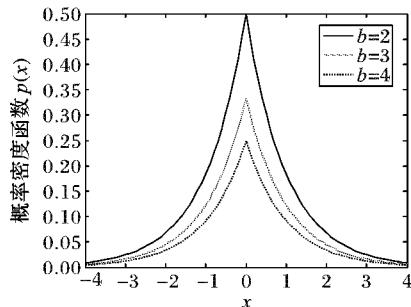


图1 拉普拉斯概率密度函数

Fig. 1 Laplace probability density function

假设数据集为 D ，查询函数是 F ，查询结果是 $F(D)$ ，添加噪声的函数为 W ，其响应值定义如下：

$$W(D) = F(D) + \text{Laplace}(\Delta F/\varepsilon) \quad (12)$$

则称 $W(D)$ 满足 ε - 差分隐私保护。

1.3.2 指数机制

Laplace 机制仅适用于查询结果是数值型的数据，而在实际的生活应用中许多数据的查询是非数值型的，因此在文献[8]中提出了一种非数值型的算法。该算法定义如下：设随机算法 F ，查询结果集为 R ， t 是 R 中某一个体，在这种机制下函数 $u(D, t)$ 中 t 是所有输出项的选中项。若算法 F 以 $\exp((\varepsilon u(D, t)) / (2\Delta u))$ 的概率从查询结果集中选择并输出的是 t ，那么称此算法满足 ε - 差分隐私的保护模型定义。

由此可以看出，当 ε 越大，被选出输出的概率就越大；当 ε 较小时，选中输出的概率就变小。

2 差分隐私保护的谱聚类算法分析

2.1 算法描述

与传统的 k -means 相比，基于图分割理论的谱聚类算法的适用性更强，不容易陷入局部最优解，能够对社交网络中非凸型的数据进行聚类。该算法把数据点看成一个个顶点，然后利用相似矩阵将样本点链接一起，采用子图最优的分割理论来划分。其中相似矩阵计算的是两个样本点之间的权重值，而这种权重可以理解为样本点的亲密度。对于数据的发布者来看，如何在发布数据的同时不泄露它们的亲密度关系是本文的研究重点。因此本文采用差分隐私结合谱聚类的方法，在相似性矩阵上加上满足拉普拉斯分布的随机噪声来达到隐私保护的目的。

基于差分隐私的谱聚类算法流程如下：假设 $p = \{p_1, p_2, \dots, p_n\}$ 和 $q = \{q_1, q_2, \dots, q_n\}$ 是 n 维空间中的两个样本集，

谱聚类算法的原理就是利用相似性函数计算样本数据集的相似性，值越大相似性越高，聚为一类的可能性就越大。同时这种相似性也可以理解为社交网络中的关系亲密度，为了确保这种亲密关系不被泄露，在计算相似性时加上差分隐私的拉普拉斯噪声来隐藏潜在的数据信息，从而实现隐私安全的保护。

算法2 基于差分隐私的谱聚类算法。

```

输入 UCI 数据集 data。
输出 标签 label。
1) 定义聚类种类 k, 根据给定的 label 计算出 k 的值;
2) 初始化 k_near = 100 和 δ = 0.9;
3) for i = 1 to row
4)   for j = 1 to row
5)     if i ≠ j
6)       dis_matij ← exp(-||pi - qj||2 / (2σ2));
7)       dis_matii ← 0 ;
8)     end if
9)   end for
10) end for
11) 保留 k_near 的权重值 Affinity[i, j]，根据积累分布函数
    (11)生成随机噪声，并添加进去；
12) 计算度矩阵和拉普拉斯矩阵的值 du[i, j], laplas[i, j];
13) 求拉普拉斯矩阵的前 k 大特征值和对应的特征向量;
14) 标准化特征向量，得到 vij = Vij / √(sum_i Vij2);
15) 利用 k-means 进行聚类，得到聚类后的 label 值;
```

2.2 算法分析

此算法在邻近数据集 D 和 D' 上添加或者删除任意一条记录时，对于每一维的数据敏感度是 1，所以全局敏感度是 n 。因为该算法是对于每个查询函数添加噪声，满足差分隐私的串行组合原理，所以这些算法组合起来的总的 $\sum_{i=1}^n \varepsilon_i$ 满足 ε - 差分隐私模型的定义衡量标准。

3 实验

3.1 实验数据集

本文采用来自于 UCI Knowledge Discovery Archive database (<http://archive.ics.uci.edu/>) 数据集中的 liver、pima、sonar、balance 四个数据集进行实验。各数据集信息如表1所示。

表1 UCI 数据集

Tab. 1 UCI datasets

数据集	样本数	属性数	种类数
liver	345	6	2
pima	768	8	2
sonar	208	60	2
balance	625	4	3

本文首先对四个数据集 liver、pima、sonar、balance 进行预处理，使其每个属性值都在区间 $[0, 1]$ ，对四个数据集分别进行谱聚类算法和差分隐私聚类算法实验，因为实验的偶然性，所以进行 20 次实验，对比 20 次实验结果的平均值。然后调整相似性函数 σ 的值，如 0.1、0.5、0.9、1.2、4、6、8、10 和 12 来确定最佳的聚类状态，并用精确度作为聚类结果的输出。从



图 2 可以看出, 聚类效果比较好的 σ 维持在 1 ~ 2。

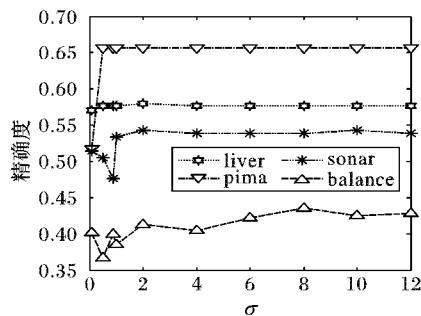


图 2 参数 σ 对聚类结果的影响

Fig. 2 Effect of parameter σ on clustering results

3.2 实验配置环境

主要采用 Matlab 软件编程来实现本文算法。实验的软硬件环境如下:

1) 硬件环境配置: Intel i5 处理器, 4 GB 内存。

2) 软件环境配置: Matlab R2013b 编程软件, 操作系统 Windows 7 64 位旗舰版。

3.3 实验结果与分析

根据前文的实验设置, 本文完成了四个数据集上的对比实验, 图 3~6 分别给出了四个数据集扰动前后的实验对比情况。

由图 3 可知, 对于数据集 liver, 运用差分隐私的谱聚类算法和只用谱聚类算法在聚类效果上是差不多的, 说明本文算法可以在具有隐私保护的前提下, 保证了数据集 liver 的聚类有效性。

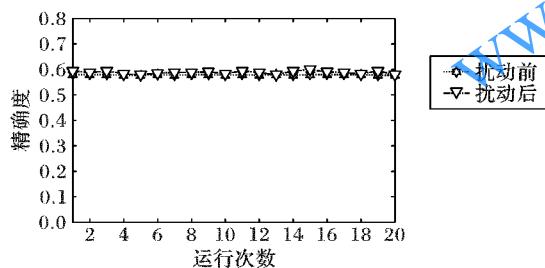


图 3 liver 数据集扰动前后的精确度结果比较

Fig. 3 Accuracy results comparison of dataset liver

由图 4 可知, 对于数据集 pima, 精确度平均分布在 0.6 ~ 0.7, 分布较为稳定, 而扰动前后的对比, 虽然总体上是加扰动前聚类效果要好, 但是此条件下并不能满足隐私保护, 所以扰动后的算法仍然具有可用性。

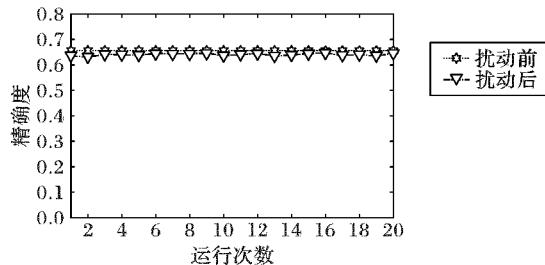


图 4 pima 数据集扰动前后的精确度结果比较

Fig. 4 Accuracy results comparison of dataset pima

由图 5 可知, 对于数据集 sonar, 其运行的总体情况是加入拉普拉斯噪声要比不加噪声好, 精确度平均分布在 0.5 ~ 0.6, 而干扰后的算法在隐私保护的前提下可以达到聚

类效果的最好状态。

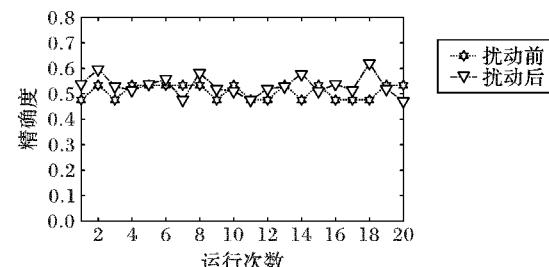


图 5 sonar 数据集扰动前后的精确度结果比较

Fig. 5 Accuracy results comparison of dataset sonar

由图 6 可知, 对于数据集 balance 的运行结果总体都比未扰动的效果更好, 其精确度的平均值稳定在 0.4 左右, 而加入扰动后的值平均在 0.5 左右, 提高了聚类的有效性。同时, 经过扰动后的权重因随机点的选取, 可能会出现样本点更好的聚类在样本中心点附近, 所以出现扰动后结果优于扰动前。

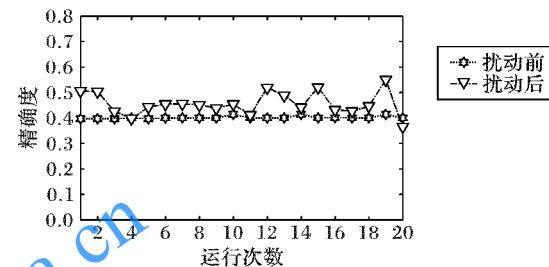


图 6 balance 数据集扰动前后的精确度结果比较

Fig. 6 Accuracy results comparison of dataset balance

对比图 3~6, 对于不同的四个数据集, 在与谱聚类算法和差分隐私聚类算法对比实验中, 扰动前的精确度总体比扰动后的结果要高一些。总体来说, 本文算法在实现隐私保护方面具有较好的成效, 同时得到的聚类精度也较高。

4 结语

本文针对传统的聚类算法存在隐私泄露的风险, 以及几个经典的聚类保护算法, 如 k -means、DBScan、 k -medoids 等, 在社交网络应用中还存在的一些不足, 利用谱聚类算法对处理凸型的空间数据和高维度的数据不容易陷入局部最优解的优势, 开展具有隐私保护的谱聚类算法研究。首先计算样本数据集之间的样本相似性作为数据点之间的权重值, 再利用差分隐私算法对权重值添加拉普拉斯分布的随机噪声, 通过干扰权重值达到隐私保护的目的。实验结果表明, 干扰后的数据不仅可以实现隐私保护, 还保证了聚类的有效性。

参考文献(References)

- [1] SHMUEL E, TASSA T. Privacy by diversity in sequential releases of databases[J]. Information Sciences, 2015, 298: 344 ~ 372.
- [2] NI W, GU M, CHEN X. Location privacy-preserving k nearest neighbor query under user's preference[J]. Knowledge-Based Systems, 2016, 103(1): 19 ~ 27.
- [3] TRUJILLO-RASUA R, DOMINGO-FERRER J. On the privacy offered by (k, δ) -anonymity[J]. Information Systems, 2013, 38(4): 491 ~ 494.
- [4] 张啸剑, 孟小峰. 面向数据发布和分析的差分隐私保护[J]. 计算机学报, 2014, 37(4): 927 ~ 949. (ZHANG X J, MENG X F. Differential privacy in data publication and analysis[J]. Chinese Journal of Computer Science, 2014, 37(4): 927 ~ 949.)



- Journal of Computers, 2014, 37(4): 927 – 949.)
- [5] OLIVEIRA S R M, ZAIANE O R. Achieving privacy preservation when sharing data for clustering[M]// JONKER W, PETKOVIC M. Secure Data Management. Berlin: Springer, 2004: 67 – 82.
- [6] MUKHERJEE S, CHEN Z, GANGOPADHYAY A. A privacy-preserving technique for Euclidean distance-based mining algorithms using Fourier-related transforms[J]. VLDB Journal, 2006, 15(4): 293 – 315.
- [7] BLUM A, DWORK C, McSHERRY F, et al. Practical privacy: the SuLQ framework[C]// Proceedings of the Twenty-Fourth ACM Sigmod-Sigact-Sigart Symposium on Principles of Database Systems. New York: ACM, 2005: 128 – 138.
- [8] DWORK C, NAOR M, PITASSI T, et al. Pan-private streaming algorithms[EB/OL]. [2018-01-10]. <http://nebula.wsimg.com/e2c5b9c40e7ca5ee436f9cb470b3ea7b?AccessKeyId=0EF19C92671ED94CE585&disposition=0&alloworigin=1>.
- [9] 李杨, 郝志峰, 温雯, 等. 差分隐私保护 k -means 聚类方法研究[J]. 计算机科学, 2013, 40(3): 287 – 290. (LI Y, HAO Z F, WEN W, et al. Research on differential privacy preserving k -means clustering[J]. Computer Science, 2013, 40(3): 287 – 290.)
- [10] 李洪成, 吴晓平, 陈燕. MapReduce 框架下支持差分隐私保护的 k -means 聚类方法[J]. 通信学报, 2016, 37(2): 124 – 130. (LI H C, WU X P, CHEN Y. k -means clustering method preserving differential privacy in MapReduce framework[J]. Journal on Communications, 2016, 37(2): 124 – 130.)
- [11] 吴伟民, 黄焕坤. 基于差分隐私保护的 DP-DBScan 聚类算法研究[J]. 计算机工程与科学, 2015, 37(4): 830 – 834. (WU W M, HUANG H K. A DP-DBScan clustering algorithm based on differential privacy preserving[J]. Computer Engineering and Science, 2015, 37(4): 830 – 834.)
- [12] 刘晓迁, 李千目. 基于聚类匿名化的差分隐私保护数据发布方法[J]. 通信学报, 2016, 37(5): 125 – 129. (LIU X Q, LI Q M. Differentially private data release based on clustering anonymization [J]. Journal on Communications, 2016, 37(5): 125 – 129.)
- [13] MATKOVIC Y, MATKOVIC Y. Robust spectral clustering for noisy data: modeling sparse corruptions improves latent embeddings[C]// Proceedings of the 2017 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2017: 737 – 746.
- [14] 熊平, 朱天清, 王晓峰. 差分隐私保护及其应用[J]. 计算机学报, 2014, 37(1): 101 – 122. (XIONG P, ZHU T Q, WANG X F. A Survey on differential privacy and applications[J]. Chinese Journal of Computers, 2014, 37(1): 101 – 122.)
- [15] DWORK C. Differential privacy: a survey of results[C]// Proceedings of the 2008 International Conference on Theory and Applications of Models of Computation. Berlin: Springer, 2008: 1 – 19.
- [16] McSHERRY F, TALWAR K. Mechanism design via differential privacy[C]// Proceedings of the 2007 IEEE Symposium on Foundations of Computer Science. Piscataway, NJ: IEEE, 2007: 94 – 103.

This work is partially supported by the National Natural Science Foundation of China (61772034, 61602009), the Natural Science Foundation of Anhui Province (1808085MF172).

ZHENG Xiaoyao, born in 1981, Ph. D. candidate, associate professor. His research interests include information security, personalized recommendation.

CHEN Dongmei, born in 1994, M. S. candidate. Her research interests include information security, intelligent computing.

LIU Yuqing, born in 1994, M. S. candidate. Her research interests include information security, data mining.

YOU Hao, born in 1997. His research interests include information security, personalized recommendation.

WANG Xiangshun, born in 1992, M. S. candidate. His research interests include personalized recommendation.

SUN Liping, born in 1980, Ph. D., professor. Her research interests include spatial data processing, intelligent computing.

(上接第 2902 页)

- [11] FENG L P, SONG L P, ZHAO Q S, et al. Modelling and stability analysis of worm propagation in wireless sensor networks [J]. Mathematical Problems in Engineering, 2015, 2015: Article ID 129598.
- [12] ZHU Q Y, CEN C. A novel computer virus propagation model under security classification [J]. Discrete Dynamics in Nature and Society, 2017, 2017: Article ID 8609082.
- [13] 孙文君, 苏旸, 曹镇. 一种非对称信息条件下的 APT 攻防博弈模型[J]. 计算机应用, 2017, 37(9): 2557 – 2562. (SONG W J, SU Y, CAO Z. Attack-defense game model for advanced persistent threats with asymmetric information [J]. Journal of Computer Applications, 2017, 37(9): 2557 – 2562.)
- [14] 唐赞玉, 刘虹. 多阶段大规模网络攻击下的网络安全态势评估方法研究[J]. 计算机科学, 2018, 45(1): 245 – 248. (TANG Z Y, LIU H. Study on evolution method of network security situation under multi-stage large-scale network attack [J]. Computer Science, 2018, 45(1): 245 – 248.)
- [15] LIU W P, ZHONG S M. Web malware spread modelling and optimal control strategies [J]. Scientific Report, 2017, 7: Article ID

42308.

- [16] 马知恩, 周义仓. 传染病动力学的数学建模与研究[M]. 北京: 科学出版社, 2004. (MA Z E, ZHOU Y C. Mathematical Modelling and Study on Infectious Disease Dynamics [M]. Beijing: Sceince Press, 2004.)

This work is partially supported by the National Natural Science Foundation of China (61503050), the Key Disciplines Construction Project of Xinzhou Teachers University (XK201403).

FENG Liping, born in 1976, Ph. D., professor. Her research interests include network security, dynamical system.

HAN Xie, born in 1964, Ph. D., professor. Her research interests include virtual reality, network security.

HAN Qi, born in 1981, Ph. D., associate professor. His research interests include network security, optimistic control, cellular neural network.

ZHENG Fang, born in 1982, M. S. lecturer. Her research interests include information security, privacy protection.